ELSEVIER

# Molecular analysis of an occlusion body protein from *Culex nigripalpus* nucleopolyhedrovirus (CuniNPV)

Omaththage P. Perera [a,*], Steven M. Valles [b], Terry B. Green [b], Susan White [b], Charles A. Strong [b], James J. Becnel [b]

[a] *Southern Insect Management Research Unit, USDA-ARS, 141 Experiment Station Road, P.O. Box 346, Stoneville, MS 38776, USA*
[b] *Center for Medical, Agricultural and Veterinary Entomology, USDA-ARS, P.O. Box 14565, Gainesville, FL 32604, USA*

## Abstract

CUN085 is an occlusion body (OB) protein from the nucleopolyhedrovirus of *Culex nigripalpus* (CuniNPV). SDS–PAGE analysis indicated that the CuniNPV OB protein is about 3 times the size (∼90 kDa) of characterized nucleopolyhedrovirus (NPVs) and granulovirus OB proteins. Rapid amplification of cDNA ends (RACE), RNase protection assay, real-time PCR, and protein sequencing were used to characterize CUN085 from CuniNPV. RACE data revealed that the transcriptional start and termination sites for the CUN085 gene yielded a polypeptide comprised of 822 amino acids indicating that translation initiates within a larger 882 amino acid open reading frame that was originally predicted from the CuniNPV genome sequence. Transcription of CUN085 started at a consensus baculovirus late transcription start site TAAG at nucleotide position 75433 of the CuniNPV genome sequence. RNase protection assays and quantitative real-time PCR show that the CUN085 transcript is first detected in mosquito larvae at approximately 6 h after infection with CuniNPV and its prevalence increased progressively over the subsequent 18 h.
Published by Elsevier Inc.

*Keywords:* CuniNPV; Nucleopolyhedrovirus of Culex nigripalpus; NPV; Occlusion body protein; CUN085; Baculovirus; Temporal expression

## 1. Introduction

The nucleopolyhedrovirus of *Culex nigripalpus* (CuniNPV) is a baculovirus that infects *Culex* species (Becnel et al., 2001). Morphological data indicates that CuniNPV is an occluded virus with amorphous, non-polyhedron-shaped occlusion bodies (Fig. 1) containing four to six individually enveloped virions (Moser et al., 2001). Although all described nucleopolyhedroviruses (NPVs) or granuloviruses (GVs) encode highly conserved occlusion body proteins called polyhedrin or granulin with molecular weights of ±29 kDa (Ahrens et al., 1997), the genome of CuniNPV appeared unique in that it did not contain an obvious gene coding for a polyhedrin-like polypeptide that fits this paradigm (Afonso et al., 2001). N-terminal amino acid sequence of an ∼30 kDa protein band isolated from denatured occlusion body proteins of CuniNPV matched the amino acid positions 691–707 of an 882 residue ORF of the CuniNPV genome. Interestingly, N-terminal analysis of the most abundant protein band (∼90 kDa) from SDS–PAGE of OB protein yielded sequence corresponding to the same 882 residue ORF beginning at the 61st residue (Afonso et al., 2001). This ORF, CUN085, spans the nucleotide position 75563–72917 on the complementary strand of the CuniNPV genome. In addition to being more than three times larger than the occlusion body proteins of other baculoviruses, the amino acid sequence of CUN085 showed no significant similarity to any known protein. Based on the amino acid sequence data of the 30 kDa protein band, it was suggested that CUN085 may code for a polypeptide that was subsequently cleaved to produce the major occlusion body protein (Afonso et al., 2001; Moser et al., 2001).

---

* Corresponding author. Fax: +1 662 686 5421.
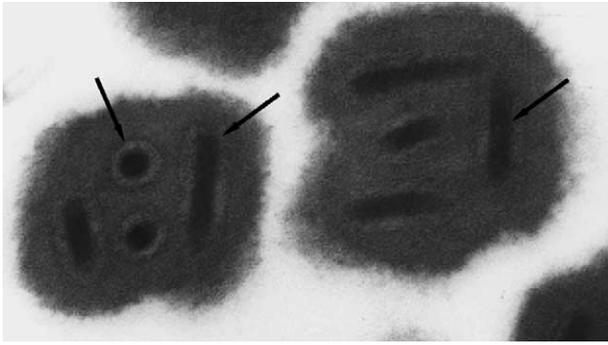 *E-mail address:* operera@msa-stoneville.ars.usda.gov (O.P. Perera).

Fig. 1. Ultrastructure of CuniNPV occlusion bodies. High magnification electron micrograph of CuniNPV virus particles. The rod shaped structures (arrows) are the infectious particles (virions) of the virus containing the viral genomic DNA. These particles are taken into the stomach of larval mosquitoes during feeding and release the virions that are responsible for infecting the midgut cells and causing larval death.

However, experimental evidence was necessary to determine if CUN085 is indeed the gene coding for the major occlusion body protein or if the CUN085 peptide is cleaved to produce the major component of the occlusion body. Therefore, we set out to identify and characterize the major occlusion body protein from CuniNPV, and analyze its transcription during the course of infection in the mosquito, *Culex nigripalpus.*

## 2. Materials and methods

### 2.1. Mosquito rearing

*Culex quinquefasciatus* eggs were collected from the mosquito rearing facility at Center for Medical, Agricultural, and Veterinary Entomology, USDA-ARS, Gainesville, Florida. Eggs were hatched in 250 ml plastic cups and approximately 3000 first instar larvae were transferred to each plastic rearing tray (45 × 90 cm) containing 3000 ml of deionized water, 2 g of larval food (2:1 hogchow:alfalfa slurry). All larvae were reared in chambers set at 28 °C and 12:12 h light:dark cycles.

### 2.2. Virus preparations

Viral occlusion body preparations were carried out as described in Moser et al. (2001). Briefly, late second instar *C. quinquefasciatus* larvae were exposed to CuniNPV by inoculating rearing trays with infected larvae or purified occlusion bodies. Larvae were collected 48 h after inoculation by straining through a nylon mesh and rinsing with deionized water to remove food particles. Collected larvae were homogenized in a glass jar using a blender. The homogenate was filtered through a 5 micron low-protein binding Durapore disc (Millipore) to eliminate cuticular debris. The filtrate was layered on top of a 30% HS-40 Ludox (Dupont) gradient and centrifuged at 15,000g for 30 min in a Sorvall high speed centrifuge (Model RC-5B) using a swinging bucket rotor. The occlusion body band

formed at a density of approximately 1.14–1.18 g/ml was carefully removed using a pasteur pipette and washed twice with 0.1 mM NaOH solution followed by pelleting with centrifugation at 10,000g.

### 2.3. Polyacrylamide gel electrophoresis

Purified occlusion bodies were denatured by re-suspension in 10 mM NaOH, 8 M Urea, or heating (100 °C) in sample loading buffer containing β-mercaptoethanol (BME) for 5 min. Denatured occlusion bodies were separated by SDS–PAGE on a 12% running gel and 5% stacking gel (Laemmli, 1970) at 25 v for 15 min followed by 100 v for 1 h using the Protean Mini gel system (Bio-Rad, Hercules, CA). Protein bands were detected by washing the gels for 10 min in distilled water and staining with Gelcode protein stain (Pierce, Rockford, IL) for 1 h. Gels were destained with two changes of distilled water for 30 min each.

Western blots for peptide sequencing were made by equilibrating, freshly run SDS–PAGE protein gels in transfer buffer (10 mM NaHCO$_3$, 3 mM Na$_2$CO$_3$, 20% Methanol, and 0.1% SDS, pH = 9.9) for 15 min and electro blotting onto polyvinylidene difluoride (PVDF) membranes (Immobilon-P$^{sq}$, Millipore, Bedford, MA) using the Transblot mini apparatus (Bio-Rad). Electroblotting was carried out at constant amperage of 350 mA for 2 h. PVDF membranes were stained with Coomassie blue R250 (Bio-Rad) using standard protocols.

### 2.4. DNA and protein sequence analysis

Dideoxy chain termination sequencing of DNA and Edman degradation sequencing of polypeptides were carried out at the core facilities for DNA and protein sequencing, respectively, at the Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, Florida. DNA and protein data analyses were carried out using the Mac Vector 7.0 (Accelrys, PA) software package and online databases such as GenBank, ProDom, ProCite, and SignalP.

### 2.5. Plasmids and competent cells

The T-A cloning vector pGEM T Easy plasmid (Promega, Madison, WI) was used in cloning all PCR amplified products. Ligations were performed using the manufacturer's suggested protocols with T4 DNA ligase and rapid ligase buffer provided with the kit. DH5α competent cells were transformed with ligations and grown on LB media plates supplemented with Ampicillin and X-Gal.

### 2.6. Determination of transcription start and termination sites

Late second-instar larvae of *Culex quinquefasciatus* were infected with CuniNPV and infected larvae were harvested at 6–12 h time intervals for 48 h. At the end of the 48 h period, all the larvae were pooled together and total RNA

was extracted using TriZol reagent (Invitrogen, Carlsbad, CA). Messenger RNA was purified from total RNA samples using the PolyATtract (Promega) mRNA purification system. A total of 5 μg of purified mRNA was used in the synthesis of cDNA from full-length mRNA using the Gene Racer kit (Invitrogen) following the manufacturer's protocol. Briefly, mRNA was subjected to a dephosphorylation step to remove the 5′-phosphate group from degraded mRNA molecules followed by a decapping step to remove the 5′-cap structure to expose the phosphate group from full-length mRNA. Then, an RNA oligonucleotide anchor was ligated to the 5-end of full-length mRNA and cDNA was synthesized using an Oligo(dT)-Anchor primer. A primer specific to the 5′- or 3′ anchor region and a gene specific primer can be used to amplify full-length 5′- or 3′-ends, respectively, of a given gene. PCR amplifications were performed with 29CnNPV72723F (5′-ACAGGATCC GTTGAATTTCTTGCCCTCTACA-3′) as the gene specific primer and the 5′-anchor primer-25RACE-5′end (5′AAGGAATTCCGACTGGAGCACGAGGACACTG A3′) to amplify 5′end of the CUN085 mRNA. The 3′ anchor primer 26RACE-3′-end (5′CATCTCGAGTGTC AACGATACGCTACGTAACG-3′) and the gene specific primer 30CnNPV75668 (5′-AGTACTAGTAACTTTTAA CGCTATGAATGTGCAA-3′) were used to amplify the 3′ end of the CUN085 mRNA. PCR products were purified by agarose gel electrophoresis, cloned into pGEM-T Easy vector, and recombinant colonies were selected for DNA sequence determination.

### 2.7. Ribonuclease protection assays

Forward primer 195CUN85-RPA-F (5′-AAACCAAC ACCTGATTCGGAG-3′) and reverse primer 196CUN85-RPA-R (5′-AAGGACTTTTGCGTGGAGAC-3′) were used to amplify a 252 bp fragment of CUN085 gene (between nucleotide positions 73773 and 74024 reference to the CuniNPV genome). DNA sequences of cloned PCR fragments were determined and a clone (pCun85RPA-3) with nucleotide sequences identical to published CuniNPV sequence was selected as the riboprobe. Similarly, a 571 bp fragment of *White* gene from *Culex quinquefasciatus* was amplified using the primers 41CQWhite F (5′-TAAGT TAACACGATCAAGAAGATCTGCGACTC-3′) and 42CqWhite R (5′-ATACCCGGGGTGATCGAGGTGA ACACGAA CG-3′), and cloned into pGem T Easy vector. A clone (pCxWhiteRPA-1) identical to published *white* gene *C. quinquefasciatus* (Accession No. U73831, Besansky and Fahey, 1997) was selected as the positive control probe. To produce antisense RNA transcripts, the pCun85RPA-3 clone was linearized at the *Nsi* I of the pGemT Easy vector. T7 RNA polymerase (MaxiScript T7 kit; Ambion, Austin, Texas) was used to synthesize a 350 nt antisense probe from the linearized Cun85RPA-3 clone. The positive control clone, pCxWhiteRPA-1, was truncated at an internal *Bam*HI restriction site and *Nco* I site in the vector, ends were filled with Klenow fragment of DNA

polymerase I (New England Biolabs, Ipswich, MA)and re-ligated to produce an insert of 205 bp. Resulting clone was linearized with *Eco*RI and antisense RNA was synthesized with T7 RNA polymerase. Each 20 μl transcription reaction contained lug of linearized plasmid, 10 mM ATP, GTP and TTP, 1 mM CTP, 5 μl of $^{32}$P-CTP (3000 Ci/mM; Perkin-Elmer, La Jolla, CA), and 1 μl of T7 RNA polymerase. RNA Millennium Marker templates (Ambion, Austin, TX) were transcribed to produce RNA size standards using same reaction conditions as above, except that 10 mM CTP and 1 μl of $^{32}$P-CTP was used to reduce the specific activity of the RNA standards. All transcription reactions were incubated at 37 °C for one hour, followed by incubation with 1 μl of RNase free DNase (Ambion, Austin, TX) for 15 min at 37 °C to remove plasmid templates from the reaction. Radio-labeled RNA transcripts (riboprobes) produced in these reactions were electrophoretically separated on a 4% polyacrylamide gel containing 8 M urea with 1× TBE (89 mM Tris base, 89 mM boric acid, 2 mM EDTA, pH 8.0) as the running buffer. The gel was wrapped with plastic wrap and an X-ray film was exposed for 5 min to identify locations of the maximum length probe bands. The autoradiograph was aligned with the gel and probe bands were cut out of the gel using sterile razor blades. The poly-acrylamide gel bands containing the full-length riboprobes were incubated overnight in probe elution buffer from the RPA III kit (Ambion, Austin, TX) to elute probe from the gel.

Approximately 3000 late second instar larvae of *Culex quinquefasciatus* were exposed to freshly purified occlusion bodies of CuniNPV in a plastic tray (45 × 90 cm) containing 3000 ml of 10 mM MgSO$_4$ and 10 ml of larval food. An estimated 10$^9$ occlusion bodies were added to the container at the beginning of the experiment. The larvae were allowed to feed for 2 h and were washed with distilled water several times to remove the CuniNPV occlusion bodies. Then, the larvae were transferred to a fresh tray of the same size containing 3000 ml of distilled water and larval food. Approximately 150 larvae were removed from the tray every hour for 18 h, beginning at the second hour after infection and immediately frozen on dry ice. Another sample of about 100 larvae was collected at the 24th hour after infection and the remainder was kept for assessing the infection rate. These larvae were visually examined under a dissecting microscope to calculate the percentage of larvae showing identifiable infections at 48 h after infection. All the samples were eventually transferred from dry ice to a −80 °C freezer for storage. The experiment was repeated five times and the two experiments showing the highest infection rates (50–90%) were selected for profiling the expression of the occlusion body gene, CUN085.

Total RNA was isolated from the samples using TriZol Reagent (Invitrogen). The pellets were dried and resuspended in 100 μL of RNase free water. The RNA samples were incubated for 30 min with 1 μl of DNAse (Ambion, Austin, TX) to remove any contaminating DNA and the RNA was precipitated by adding 5 μL of 5 M sodium

acetate and 250 μL of 100% ethanol. RNA samples were quantified by spectrophotometry.

Five micrograms of total RNA isolated from larvae was mixed with 10,000 cpm of the CUN085 RPA probe and CxWhite RPA probe. RNase protection assays were performed using manufacturer's protocol for RPA III kit (Ambion, Austin, TX). Protected RNA fragments were resolved on a 4% polyacrylamide gel containing 8 M urea using 1× TBE running buffer. At the end of the run, the gel apparatus was disassembled, the gel rinsed with 10% acetic acid to remove urea, and then transferred to a filter paper. The gels were dried using a heated vacuum dryer and autoradiography was performed to visualize the bands.

### 2.8. Real-time quantitative PCR

Approximately 3000 late second instar larvae of *Culex quinquefasciatus* were exposed to freshly purified occlusion bodies of CuniNPV in a plastic tray containing 3000 ml of deionized water and 2 g larval food. Next, 21 mL of 2 M Epson salt (MgSO$_4$) and 10 ml of stock virus (final concentration $1 \times 10^8$ OBs/mL in 14 mM Epson salt) were added to the deionized water. The larvae were allowed to feed for 4 h before 21 mL of 2 M CaCl$_2$ was added to the tray to stop transmission (Becnel et al., 2001). Approximately 200 larvae were removed from the tray at 2, 4, 6, 8, 10, 12, 18, 24, and 48 h, beginning at the second hour after infection. The larvae were strained through a nylon mesh and rinsed with deionized water to remove food particles. After each time series, the samples were placed in a −80 °C freezer for storage. Approximately 200 larvae were visually examined under a dissecting microscope to calculate the percentage showing patent infections after 48 h. The experiment was repeated twice with an average infection rate of (91.8% ± 2.3%, mean ± SE). Controls, consisting of 100 or 200 larvae, were removed from the tray prior to treating and held in cups with strained food in groups of 100 or 200 larvae in 100 ml of deionized water.

Total RNA was isolated from the samples using TriZol Reagent (Invitrogen). RNA samples were incubated for 30 min with 3 μL of DNase (Ambion, Austin, TX) to remove any contaminating DNA, heated to 65 °C for 15 min to kill DNase activity, and the RNA was precipitated by adding 10 μL of 3 M sodium acetate and 250 μl of 100% ethanol. This procedure was repeated twice before quantifying the RNA samples with a spectrophotometer. Quantified RNA samples were used to prepare 100 ng/μl RNA solutions for use in cDNA synthesis. Five microliters (500 ng) of RNA was used in cDNA synthesis with SuperScript II First Strand cDNA synthesis kit (Invitrogen). Each 20 μl reverse transcription reaction contained 4 μl of 5× first strand synthesis buffer, 1 μl of 10 mM dNTP mix, 1 μl of 50 nM random hexamer mix, 1 μl of 100 mM Dithiothrietol (DTT), and 200 units of Superscript II reverse transcriptase. After addition of reverse transcriptase, reactions were mixed well, briefly centrifuged, and incubated at room temperature for 10 min followed by 50 min incuba-

tion at 42 °C. Reverse transcription reactions were terminated by incubating samples at 70 °C for 15 min and stored at −20 °C until used.

CuniNPV genomic DNA, extracted from purified occlusion bodies by Proteinase K digestion, phenol–chloroform, followed by ethanol precipitation was serially diluted 10- and 100-fold and quantified by spectrophotometry. Three independent readings from the original DNA preparation and each dilution were taken and the average of all the quantifications was used to calculate the DNA quantity in the sample. A solution of CuniNPV DNA with 4 ng/μl was prepared from the quantified original DNA preparation and was used to prepare 10-fold serial dilutions containing 0.4, 0.04, 0.004, and 0.0004 ng/μl standards for quantitative PCR. 2.5 μl of each dilution was used in triplicate to yield a dilution series of 1, 0.1, 0.01, and 0.001 ng DNA.

Real-time quantitative PCR (QPCR) primers and the Taqman probe designed to detect CUN085 transcripts were purchased through Assays-By-Design service provided by Applied Biosystems, Inc. 96-well plates, optical adhesive film, and Universal PCR master mix with Urasyl-N-Glycosylase (UNG) were purchased from Applied Biosystems. Real-time quantitative PCR was performed in ABI model 7500 Sequence Detection System using the absolute quantification method. Forward primer (5′-CGGTTGTTGG CCAGAATGTAG-3′) and reverse primer (5′-TGGTTG GTGTGAACTCCATTCTAC-3′) sequences were designed to amplify 41 bp amplicon (from nucleotide position 73084 to 73124, inclusive, of the CuniNPV genome). Taqman MGB detection probe sequence was labeled at 5′ end with FAM dye (5′-FAM-CACTCCCCGTCGTTCG-3′). Primer /probe concentrate (20×) consisted of 18 μM of each primer and 5 μM of detection probe. Assay reactions were set up in 96-well reaction plates using 4 μl of cDNA (equivalent to 100 ng of total RNA) from each reverse transcription, 10 μl of 2× universal PCR master mix, 1 μl of 20× primer/probe mix and 5 μl of water to bring total reaction volume to 20 μl. Assay plates were covered with optical adhesive film and briefly spun to bring all the liquid to the bottom of the wells. Reactions were incubated at 50 °C for 2 min to activate UNG followed by 10 min incubation at 95 °C to activate *Taq* polymerase. Forty cycles of amplifications with 15 s denaturing step at 95 °C and 1 min annealing/extension step at 60 °C minutes were performed. Spectral data were acquired during the annealing/extension step and transcripts were quantified using the standard curve generated by the serial dilutions of CuniNPV genomic DNA. Each time point of the two separate experiments was analyzed in duplicate and the estimated quantities were averaged.

### 2.9. Informatics

DNA and protein sequence analyses were performed using the Mac Vector software package (Accelrys, San Diego, CA). Database searches (GenBank, EMBL, SWISS-PROT, and PIR) for DNA and protein comparisons were carried out using BLAST and PSI-BLAST

(Altschul et al., 1990, 1997; Pearson, 1990). Online software available on the ExPASy server Online software available on the ExPASy server (Gasteiger et al., 2003) and Swiss EMBNet Node (www.ch.embnet.org) wer e used for prediction of secondary structure and searching for domains, transmembrane regions and subcellular localization (Reinhardt and Hubbard, 1998).

## 3. Results and discussion

### 3.1. Identification of major occlusion body protein

Because polyhedrin proteins are the most abundant components of the occlusion bodies of nucleopolyhedroviruses, determination of the amino acid sequence of the most abundant protein band in denatured occlusion bodies have been used to identify novel occlusion body genes (Rohrmann, 1992; Rohrmann et al., 1979, 1981; Serebryani et al., 1977). SDS–PAGE of occlusion bodies (OB) denatured with 10 mM NaOH consistently showed a banding pattern characteristic of CuniNPV (Fig. 2 SDS–PAGE of CuniNPV OB). OBs denatured using 8 M urea or 5% BME and heat produced the same banding profile as with 10 mM NaOH (data not shown). Amino acid sequence of the most abundant protein (Fig. 2, arrow), approximately 90 kDa, matched the amino acid sequence beginning at the 61st amino acid (a Methionine) of predicted CUN085 open reading frame (Afonso et al., 2001). The molecular weight of band 1 corresponds to the calculated molecular weight for CUN085 polypeptide of 822 residues, starting at the 61st amino acid of the predicted ORF. In addition, amino



Fig. 2. SDS Analysis of CuniNPV OB. Purified OB (lane 2) was run on a 12% SDS–PAGE with a broad range marker (lane 1). The arrow indicates the major occlusion body protein and the asterisks indicate truncated CUN085 protein bands.

acid sequences of 2 other less abundant and smaller proteins (Fig. 2, asterisks) also matched the sequence of CUN085, beginning at the 61st amino acid. Thus we assumed that these smaller bands may have been generated by loss of amino acids from the C-terminal end of CUN085. Because none of the N-terminal amino acid sequences derived from the protein bands started at the first methionine of the predicted ORF, it was initially postulated that either the polypeptide was post-translationally cleaved between the 60th and 61st residues or the translation actually started at the methionine at the 61st codon of the predicted ORF. However, cDNA end analysis (see below) indicated that translation of CUN085 can only begin at the nucleotide position 75383 of the CuniNPV genome (Accession No. AF403738), starting with the second methionine of the predicted ORF (Fig. 3, nucleotide position 181). Therefore, we conclude that CUN085, the major OB protein of CuniNPV, is a polypeptide of 822 amino acids with a translation start at the nucleotide position 75383 of the CuniNPV genome and not a cleavage product of the 882 amino acid polypeptide translated from the predicted ORF.

### 3.2. Transcription start and termination

Nucleotide sequence analysis of cloned PCR fragments generated by 5′-RACE indicated that transcription of the major OB protein gene of CuniNPV starts at either the first or second A residue of a baculovirus late transcription promoter/start site sequence, TAAG, beginning at nucleotide position 75433 (Fig. 3, nucleotide position 132) of the CuniNPV genome sequence. Of the six clones sequenced, five initiated at the first A and only one initiated at the second A of the TAAG sequence (Fig. 3, arrows). This late transcription start sequence is located 131 bp downstream of the translation start site of the ORF predicted from the genome sequence, precluding possible production of the CUN085 polypeptide initiating at the beginning of the predicted ORF (Fig. 3).

Analysis of DNA sequences of 8 randomly selected clones isolated from 3′-RACE revealed that the majority of the CUN085 transcripts terminated at either nucleotide position 72770 or 72838 of the CuniNPV genome sequence (Fig. 3, nucleotide positions 2794 and 2728, respectively). Both these termination sites were located less than 20 nucleotides downstream of consensus eukaryotic polyadenylation signals (AATAAA). Two transcripts that terminated at nucleotide position 72885 (Fig. 3, nucleotide position 2679), a few nucleotides upstream of the first polyadenylation signal, were also discovered. It is likely that these clones resulted from mispriming of the oligo dT primer during first strand cDNA synthesis at the stretch of 14 adenosine residues located at this site. However, these shorter transcripts could also have been produced by a viral late gene polymerase which may use an alternative polyadenylation signal. It has been demonstrated that the transcription of AcMNPV late genes are carried out by virally encoded DNA-directed RNA polymerases and transcription factors which do not require eukaryotic
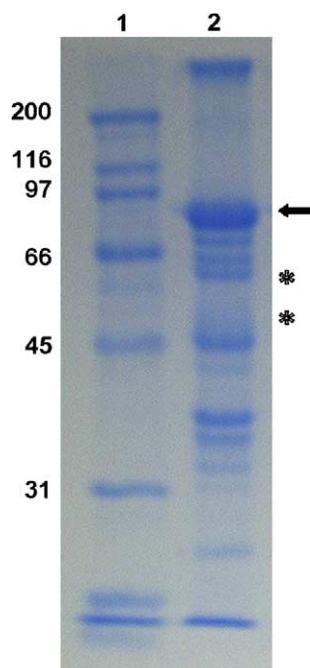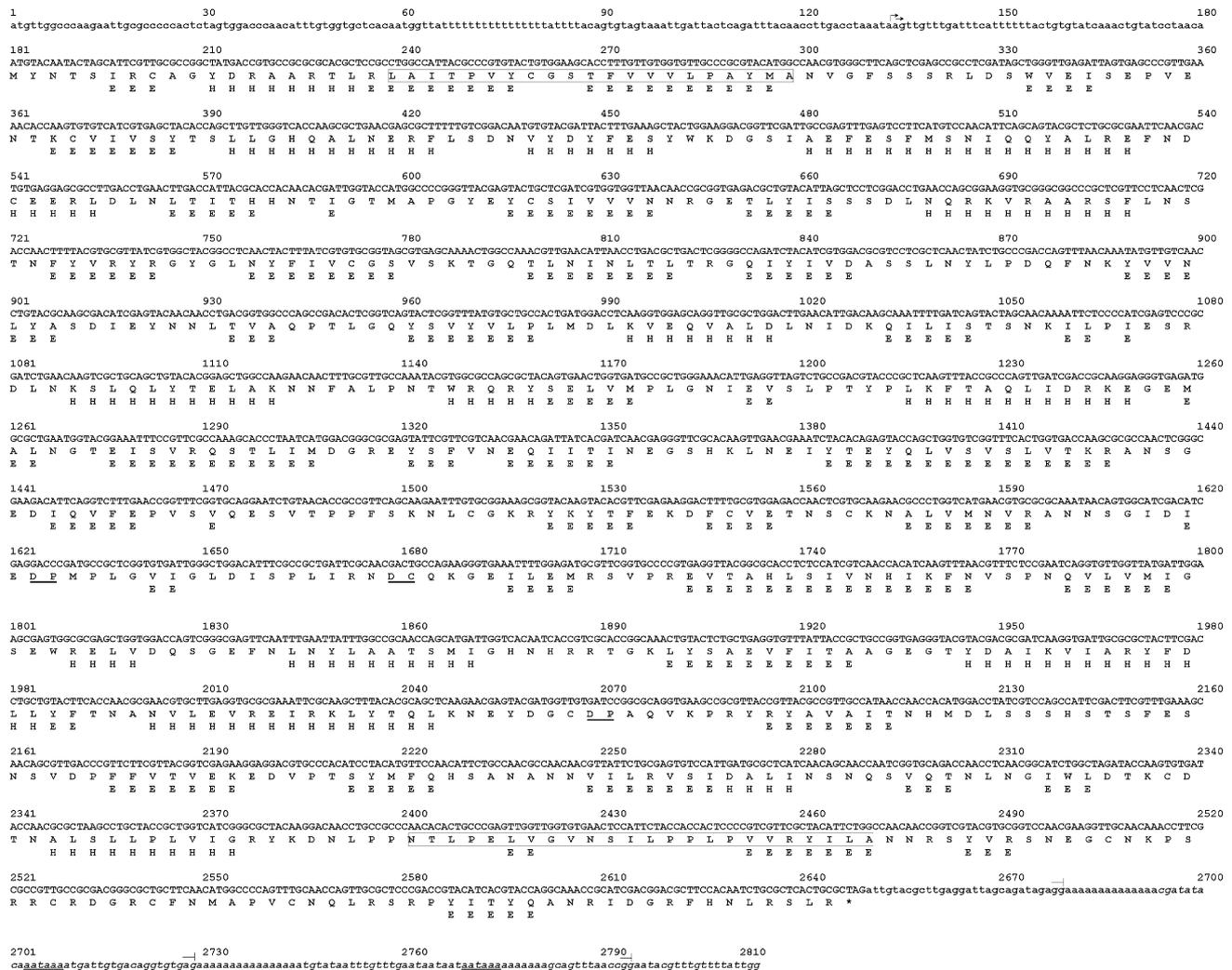
Fig. 3. Nucleotide sequence of the predicted CUN085 open reading frame (ORF) and experimentally determined ORF and acid sequences. Transcriptional start sites determined by 5′-RACE is shown by arrows and the end of the transcripts determined by 3′-RACE is marked by the symbol (⊣). Consensus eukaryotic polyadenylation signals are double underlined. Amino acid sequence based on the experimental data is given below the nucleic acid sequence. Most predominant polypeptide cleavage sites at residues 483, 501, and 631 are marked by underlined amino acids symbols. Putative transmembrane domains are boxed and secondary structure predictions of strand/extended (E) or helical (H) structures are given below the amino acid sequence. Areas without any secondary structure symbol indicate residues without any predicted structure (random coil or unstructured).

consensus polyadenylation signals for processing 3′-ends of the transcripts (Guarino et al., 1998; Jin and Guarino, 2000). The RNA polymerase purified from AcMNPV consisted of four subunits coded by p47, LEF-4, LEF-8, and LEF-9 genes (Guarino et al., 1998). Homologues of AcMNPV p47, LEF-4, LEF-8, and LEF-9 were identified in the genome of Cuni-NPV (Afonso et al., 2001). However, they exhibited only 23–26% amino acid identity to AcMNPV counterparts. Whether an alternative polyadenylation/processing mechanism is used in the transcription of the CuniNPV late genes is not known.

### 3.3. Temporal expression of CUN085 using RNase protection and real-time PCR

Two independent RNase protection assays were conducted with total RNA extracted from larvae collected in two different time-course experiments. RNA from larvae at the start of the experiment (0), 2, 4 −11, and 24 h after infection was subjected to RNase protection analysis (Fig. 4). Transcripts of the *white eye* gene of *Culex quinquefasciatus* served as the internal control for the samples. In this RNase protection assay, the transcription of CUN085 could be detected as early as 8 h after infection while the second assay indicated a start of transcription at 9 h after
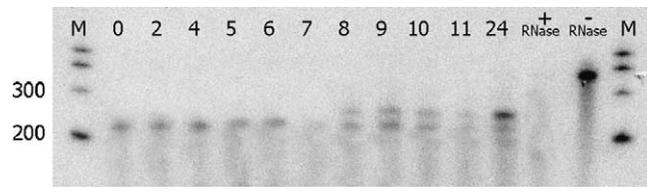


Fig. 4. Temporal expression of Cun85. An RNase protection assay of the OB gene, Cun85 was performed with RNA samples collected from infected larvae at various time points (0, 2, 4, 5, 6, 7, 8, 9, 10, 11, and 24 h). Transcription of Cun85 was detected as early as 8 h post infection. Controls with/without RNase are also indicated on the gel.

infection. This variability may be due to the differences in intensity of CuniNPV infection of the larvae between two experiments.

Real-time quantitative PCR assays performed using the RNA extracted from larvae collected in two additional independent time course experiments detected CUN085 transcripts at 6 h (Fig. 5), corroborating the RNase protection assay results of 8–9 h. Absolute quantification of the RNA from the sample 6 h after infection indicated the presence of approximately 4000 copies of CUN085 transcripts in a cDNA quantity prepared from 100 ng of total RNA. Compared with 6 h after infection, 8 hrs after infection showed over 12-fold increase in the CUN085 transcripts followed by 21-, 388-, 5644-, and 46448-fold increases in samples collected at 10, 12, 18, and 24 h, respectively, after inoculation. This rapid, steady increase of CUN085 transcripts correlates well with the progression of the infection (Moser et al., 2001) where occluded viruses may be observed under light microscope between 12–15 h after infection. No amplification was detectable in the mock infected controls, or in the samples at 0, 2, or 4 h after inoculation.

Temporal gene expression studies on baculoviruses of *Autographa californica* (AcMNPV) and *Lymantria dispar* (LdMNPV) have reported detection of polyhedrin transcripts at 6 and 24 h after infection of cultured cells, respectively (Friesen and Miller, 1985; Riegel and Slavicek, 1997). Cell cultures are usually infected at very high multiplicity of infection (m.o.i) that lead to production of detectable levels of transcripts in a reproducible manner, in contrast to low m.o.i infections through oral ingestion of occluded viruses by mosquito larvae. Although direct comparison of gene expression levels may not be possible between experiments carried out using cell cultures and mosquito larvae, the detection of CUN085 (CuniNPV occlusion body protein) transcripts at 6 h after infection indicates that the promoter

is activated at a very early stage of infection. However, polyhedrin polypeptide production has been detected in cell cultures infected with AcMNPV only 12–15 h after infection. Therefore, whether the production of CUN085 polypeptides occurs during very early stages of infection must be further investigated using protein detection methods.

BLAST searches against non-redundant databases (translated or non-translated) revealed no significant similarities to DNA or protein sequences available to date. Transmembrane domain prediction using TMpred revealed two putative transmembrane domains spanning residues 20–40 (inside-to-outside, Score = 1444) and residues 741–764 (outside-to-inside, Score = 649). Secondary structure predictions, including protein disorder and globularity prediction by DisEMBL and GlobPlot ( Linding et al., 2003a,b) predicted a globular protein with 24.2% helical, 20.3% strand, and 55.5% unstructured (random coil/loop) regions (Fig. 3). PROSITE and Pfam domain predictions did not yield any significant structural similarities to known patterns or domains. The fact that this occlusion body protein is mostly unstructured is not surprising. Previous studies of the OBs and virions at various time points (Becnel et al., 2001) show the release of ODVs within the OBs. The random coil nature of the occlusion body protein would possibly explain how ODVs transverse the OBs and are released to infect midgut cells without being degraded.

Edman degradation sequencing of protein gel bands isolated from denatured OB proteins also identified the presence of C-terminal fragments of CUN085 polypeptide cleaved at residues 483, 501, and 631 as well as N-terminal fragments that are smaller in size than the full-length CUN085 polypeptide (data not shown). These bands were consistently observed in SDS–PAGE at low abundance and were considered cleavage products of the CUN085 protein. Computer analysis of the amino acid sequence revealed formic acid cleavage sites, but no protease cut sites at the above positions. The only characteristic common among the three cleavage sites was the presence of an aspartic acid (D) residue at the P1 position (D-P or D-C). In addition, calculations based on cleavage positions and molecular weight indicated that these C-terminal fragments of CUN085 can be paired with N-terminal fragments identified by amino acid sequencing to yield the molecular weight of the full-length CUN085 protein. Because these cleavage products were consistently present as minor bands in denatured OB protein gels, it is possible that these cleavage sites were subjected to digestion by larval gut proteases during the OB isolation process.

We have utilized amino acid sequencing, RACE, RNase protection assay, and real-time PCR to identify and characterize the temporal expression of CUN085, the major occlusion body protein of CuniNPV, during the course of natural infection of Culex quinquefasciatus larvae. Experimental evidence from cDNA end determinations clearly indicated that transcription of the CUN085 gene started at nucleotide positions 75433 or 75432 and terminated at nucleotide position 72770, 72871, or 72838 of the CuniNPV
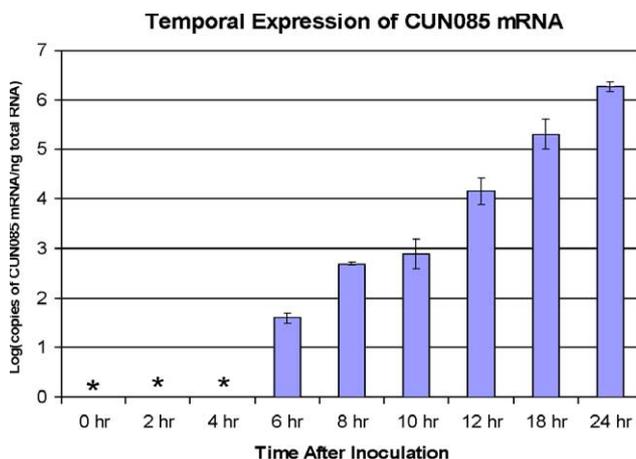


**Temporal Expression of CUN085 mRNA**

Fig. 5. Analysis of Cun85 by Quantitative real-time PCR. Log values of CUN085 mRNA quantity (copies per ng total RNA) at different time points after inoculation indicated detectable levels of CUN085 transcripts at 6 h after inoculation. Asterisks at time points 0, 2, and 4 h after inoculation indicate that CUN085 mRNA was not detected.

genome sequence. The transcription of CUN085 started at a consensus baculovirus late transcription start site (Lu et al., 1996;Morris and Miller, 1994; Ooi et al., 1989) and the majority of the transcripts terminated after one of the two consensus eukaryotic polyadenylation signals. The occlusion body protein ORF spanned nucleotide positions 75384–72912 (Fig. 2, nucleotide positions 181–2646), coding for an 822 residue polypeptide (MW: 93.2 kDa) and lacked the first 60 amino terminal amino acids present in the 882 residue molecule predicted from the CUN085 ORF described previously (Afonso et al., 2001). Translation start of CUN085 at the nucleotide position 75384 of the CuniNPV genome was confirmed by amino acid sequencing of CUN085 protein isolated from purified occlusion bodies.

The size of CUN085 is much larger than the occlusion body proteins of lepidopteran and hymenopteran baculoviruses annotated thus far and does not show sequence similarity to any protein sequence in the available databases (Garcia-Maruniak et al., 2004; Lauzon et al., 2004). It has previously been suggested that the lack of homology for the OB protein of mosquito baculoviruses with polyhedrins and granulins of terrestrial insects may be due to specific adaptations that allow them to survive and persist in the aquatic environment (Moser et al., 2001). It should also be noted that CuniNPV is a highly divergent baculovirus with only 36 genes (33.0% of the total genes) showing any amino acid sequence similarity to the genes present in other annotated baculoviruses (Afonso et al., 2001). Protein sequence analyses of the OBs and ODVs have revealed that several of the previously unknown CuniNPV genes are structural components of the OB or ODV that may perform functions similar to those of conserved baculovirus structural proteins (data not shown). Future structural studies may reveal if CUN085 has any functional similarities to occlusion body proteins of other baculoviruses.

## Acknowledgment

## References

Afonso, C.L., Tulman, E.R., Lu, Z., Balinsky, C.A., Moser, B.A., Becnel, J.J., Rock, D.L., Kutish, G.F., 2001. Genome sequence of a baculovirus pathogenic for *Culex nigripalpus*. J. Virol. 75, 11157–11165.

Ahrens, C.H., Russell, R.L.Q., Funk, C.J., Evans, J.T., Harwood, S.H., Rohrmann, G.F., 1997. The sequence of the *Orgyia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus genome. Virology 229, 381–399.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J.Mol.Biol. 215, 403–410.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acids Res. 25, 3389–3402.

Becnel, J., White, S., Moser, B., Fukuda, T., Rotstein, M., Undeen, A., Cockburn, A., 2001. Epizootiology and transmission of a newly discovered baculovirus from the mosquitoes Culex nigripalpus and C. quinquefasciatus. J. Gen. Virol. 82, 275–282.

Besansky, N.J., Fahey, G.T., 1997. Utility of the white gene in estimating phylogenetic relationships among mosquitoes (Diptera: Culicidae). Mol. Biol. Evol. 14 (4), 442–454.

Friesen, P.D., Miller, L.K., 1985. Temporal regulation of baculovirus RNA: overlapping early and late transcripts. J. Virol. 54, 392–400.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A., 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31, 3784–3788.

Guarino, L.A., Xu, B., Jin, J., Dong, W., 1998. A virus-encoded RNA polymerase purified from baculovirus-infected cells. J. Virol. 72, 7985–7991.

Jin, J., Guarino, L.A., 2000. 3′-End formation of baculovirus late RNAs. J. Virol. 74, 8930–8937.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 681–685.

Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J., Russell, R.B., 2003a. Protein disorder prediction: implications for structural proteomics. Structure 11, 1453–1459.

Linding, R., Russell, R.B., Neduva, V., Gibson, T.J., 2003b. GlobPlot: exploring protein sequences for globularity and disorder. Nucleic Acids Res. 31, 3701–3708.

Lu, A., Craig, A., Casselman, R., Carstens, E.B., 1996. Nucleotide sequence, insertional mutagenesis, and transcriptional mapping of a conserved region of the baculovirus Autographa californica nuclear polyhedrosis virus (map unit 64.8–66.9). Can. J. Microbiol. 42, 1267–1273.

Morris, T.D., Miller, L.K., 1994. Mutational analysis of a baculovirus major late promoter. Gene 140, 147–153.

Moser, B.A., Becnel, J.J., White, S.E., Alfonso, C., Kutish, G., Shanker, S., Almira, E., 2001. Morphological and molecular evidence that *Culex nigripalpus* baculovirus is an unusual member of the family *Baculoviridae*. J. Gen. Virol. 82, 283–297.

Ooi, B.G., Rankin, C., Miller, L.K., 1989. Downstream sequences augment transcription from the essential initiation site of a baculovirus polyhedrin gene. J. Mol. Biol. 210, 721–736.

Pearson, W.R., 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. Methods in Enzymology 183, 63–98.

Reinhardt, A., Hubbard, T., 1998. Using neural networks for prediction of the subcellular location of proteins. Nucleic Acids Res. 26, 2230–2236.

Riegel, C.I., Slavicek, J.M., 1997. Characterization of the replication cycle *of the Lymantria dispar* nuclear polyhedrosis virus. Virus Res. 51, 9–17.

Rohrmann, G.F., 1992. Baculovirus Structural Proteins. J. Gen. Virol. 73, 749–761.

Rohrmann, G.F., Bailey, T.J., Brimhall, B., Becker, R.R., Beaudreau, G.S., 1979. Tryptic peptide analysis and NH2-terminal amino acid sequences of polyhedrins of two baculoviruses from Orgyia pseudotsugata. Proc. Natl. Acad. Sci. USA 76, 4976–4980.

Rohrmann, G.F., Pearson, M.N., Bailey, T.J., Becker, R.R., Beaudreau, G.S., 1981. N-Terminal polyhedrin sequences and occluded Baculovirus evolution. J. Mol. Evol. 17, 323–329.

Serebryani, S.B., Levitina, T.L., Kautsman, M.L., Radavski, Y.L., Gusak, N.M., Ovander, M.N., Sucharenko, N.V., Kozlov, E.A., 1977. The primary structure of the polyhedral protein of nuclear polyhedrosis virus (NPV) of Bombyx mori. J. Invertebr. Pathol. 30, 442–443.